

Wed May 29 16:25:20 2002

us-09-676-436-3.s2lml00.rn1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 10:15:51; Search time 94.14 Seconds
(without alignments)
13020.089 Million cell updates/sec

Title: US-09-676-436-3
Perfect score: 4990
Sequence: 1 ctgaagactctcgcagatga.....ggtattattagaagctc 4990

Scoring table: IDENTITY_NDC
Gapop 10.0, Gapext 1.0

Searched: 383533 seqs, 122816752 residues
Total number of hits satisfying chosen parameters: 613726

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	26	0.5	100	US-08-400-256-4
2	26	0.5	100	US-08-975-365-4
3	25.2	0.5	99	US-08-427-097-5
4	25.2	0.5	99	US-08-878-957-5
5	24.6	0.5	71	US-08-458-423A-78
6	24.6	0.5	71	US-08-458-424B-78
7	24.6	0.5	71	US-08-973-124-78
8	24.6	0.5	71	US-08-08014-78
9	24.6	0.5	72	US-08-100-664A-5
10	24.4	0.5	100	US-08-975-365-9
11	24.4	0.5	100	US-08-400-256-9
12	24.4	0.5	97	US-08-182-175A-54
13	24.4	0.5	97	US-08-674-633A-63
14	24.4	0.5	97	PCT-US92-06412-54
15	23.8	0.5	83	US-08-672-138A-8
16	23.6	0.5	87	US-08-631-731A-8
17	23.4	0.5	77	US-07-679-052A-8
18	23.4	0.5	60	US-08-643-704A-17
19	23.4	0.5	94	US-09-344-888A-8
20	22.8	0.5	76	US-08-246-583-9
21	22.6	0.5	96	US-08-484-322-5
22	22.6	0.5	77	US-07-679-052A-10
23	22.4	0.4	90	US-08-123-702-21
24	22.4	0.4	95	US-08-332-766A-41
25	22.4	0.4	97	US-08-182-175A-48
26	22.4	0.4	97	US-08-474-633A-57
27	22.4	0.4	97	PCT-US92-06412-48

28	22.2	0.4	71	US-08-465-591A-75	Sequence 75, Appl
29	22.2	0.4	71	US-08-465-594A-75	Sequence 75, Appl
30	22.2	0.4	71	US-08-973-124-260	Sequence 260, App
31	22.2	0.4	71	PCT-US96-08014-260	Sequence 99, Appl
32	22.2	0.4	85	US-08-488-402A-99	Sequence 99, Appl
33	22.2	0.4	85	US-08-484-552A-99	Sequence 99, Appl
34	22.2	0.4	85	PCT-US96-09472-99	Sequence 3, Appl
35	22.2	0.4	87	US-08-110-286A-3	Sequence 166, App
36	22.2	0.4	90	US-09-461-697-166	Sequence 166, App
37	21.8	0.4	54	US-08-758-306-514	Sequence 5, Appl
38	21.8	0.4	79	US-08-933-983-55	Sequence 35, Appl
39	21.8	0.4	81	US-08-445-463B-10	Sequence 10, Appl
40	21.6	0.4	65	US-08-445-464C-10	Sequence 10, Appl
41	21.6	0.4	65	PCT-US94-03437-10	Sequence 10, Appl
42	21.6	0.4	63	US-08-602-093-23	Sequence 23, Appl
43	21.6	0.4	85	US-08-413-118-113	Sequence 113, App
44	21.6	0.4	93	US-08-473-446-113	Sequence 113, App
45	21.6	0.4	93		

ALIGNMENTS

RESULT 1
US-08-400-256-4/c
Sequence 4, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asger Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/400,256
APPLICATION NUMBER: US/08/400,256
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3965.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-878-9655
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-400-256-4
Query Match 0.5%; Score 26; DB 1; Length 100;
Best Local Similarity 55.6%; Pred. No. 6.1e+02;
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 4469 aagaatgaagcctagagatagacttggaataatcttaatacactacgtatgtat 4528

Db 97 AAGCTGACGACGCTAAGGCTATCGTTGACATGTTGACTTATCTGTTCTTGTAC 38
QY 4529 tattacataagacgtgctgtagagcag 4558
Db 37 CAATTGGAATACTACTGCTGTAGACGAG 8

RESULT 2

US-08-975-365-4/C

Sequence 4, Application US/08975365

Patent No. 6011007

GENERAL INFORMATION:

APPLICANT: Havelund, Svend

APPLICANT: Halstrom, John

APPLICANT: Jonassen, Ib

APPLICANT: Andersen, Asger Sloth

APPLICANT: Markussen, Jan

TITLE OF INVENTION: ACYLATED INSULIN

NUMBER OF SEQUENCES: 49

CORRESPONDENCE ADDRESS:

ADDRESS: No. 60110070 No. 6011007disk of No. 6011007th America, Inc.

STREET: 405 Lexington Avenue, 64th Floor

CITY: New York

STATE: New York

COUNTRY: United States of America

ZIP: 10174-6401

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PC-DOS/MS-DOS

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/975,365

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/400,256

FILING DATE: 03-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: Lambiris, Elias J.

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 3985,220-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-0123

TELEFAX: 212-867-9655

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 100 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-975-365-4

Query Match
Best Local Similarity 0.5%; Score 26; DB 3; Length 100;
Matches 50; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 4469 aagaatgaagccatgaagatgactgactgaataatcttcaactactgactgataa 4528
Db 97 AAGCTGACGACGCTAAGGCTATCGTTGACATGTTGACTTATCTGTTCTTGTAC 38
QY 4529 tattacataagacgtgctgtagagcag 4558
Db 37 CAATTGGAATACTACTGCTGTAGACGAG 8

RESULT 3
US-08-427-097-5
Sequence 5, Application US/08427097
Patent No. 568294

GENERAL INFORMATION:
APPLICANT: Meagher, Richard B.
APPLICANT: Sommer, Anne O.
TITLE OF INVENTION: Metal Resistance Sequences and
TITLE OF INVENTION: Transgenic Plants
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESS: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: US
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/427,097
FILING DATE: 21-APR-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Felder, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 40-94
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-427-097-5

Query Match
Best Local Similarity 0.5%; Score 25.2; DB 1; Length 99;
Matches 51; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 3026 atgctcactgaatgtagaggtgtagaggtacaaagccaggtggcgactcaagat 3085
Db 2 AAGACCCAGCTATAGTGAAGCTGTACTGCTGATTCGATGGAAGCATTTGAAGTGC 61
QY 3086 ttgattttcacaagcaatgaaccttccttat 3119
Db 62 GTGAGCATATCTCAAGCAAGCAAGCAAGTCTTATAT 95

RESULT 4
US-08-878-957-5
Sequence 5, Application US/08878957
Patent No. 5965796
GENERAL INFORMATION:
APPLICANT: Meagher, Richard B.
APPLICANT: Summers, Anne O.
APPLICANT: Rugh, Clayton L.
TITLE OF INVENTION: Metal Resistance Sequences and
TITLE OF INVENTION: Transgenic Plants
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESS: Greenlee, Winner and Sullivan, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: US
ZIP: 80303

Wed May 29 16:25:20 2002

us-09-676-436-3.sz1ml100.rn1

Page 3

```
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,957
FILING DATE: 19-JUN-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/427,097
FILING DATE: 21-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Farber, Donna M.
REGISTRATION NUMBER: 33,878
REFERENCE/DOCKET NUMBER: 40-94A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 499-8080
TELEFAX: (303) 499-8089
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 99 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "Oligonucleotide"
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-878-957-5

Query Match          0.5%; Score 25.2; DB 2; Length 99;
Best Local Similarity 54.3%; Pred. No.1e+03;
Matches 51; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

Oy 3026 atgtcctgaactaatgtgaagatgtgtgaggtacaaagccaggtggcgactcaagat 3085
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2  AAGACCCACCTAGAGTAGAGTAGAGTGTGCTGATTCGATTCGATGGAAGCATTGAAGTGC 61
Oy 3086 ttgatttctaaagaattgaactgccttat 3119
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 GTGAGCATACTCAAGCAAGCAGTTGCTTAT 95

RESULT 5
US-08-458-423A-78/C
Sequence 78, Application US/08458423A
Patent No. 5731144
GENERAL INFORMATION:
APPLICANT: PENELOPE J. TOOTHMAN
APPLICANT: STEVEN RINGQUIST
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY TGFA NUCLEIC
TITLE OF INVENTION: ACID LIGANDS AND INHIBITORS
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESS: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,423A
FILING DATE: 2-JUNE-1995
CLASSIFICATION: 536
```

```
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA: 07/536,428
APPLICATION NUMBER: 07/536,428
FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA: 07/964,624
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA: 08/117,991
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA: 07/931,473
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McLearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX 34-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3433
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-458-423A-78

Query Match          0.5%; Score 24.6; DB 1; Length 71;
Best Local Similarity 65.5%; Pred. No.1.2e+03;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Oy 3958 gaacgtgattgagtcgtattcaagcagatcacattgcgactcaagtcctcc 4012
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 56 GAAGACCTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 2

RESULT 6
US-08-458-424B-78/C
Sequence 78, Application US/08458424B
Patent No. 5731424
GENERAL INFORMATION:
APPLICANT: PENELOPE J. TOOTHMAN
APPLICANT: STEVEN RINGQUIST
APPLICANT: LARRY GOLD
TITLE OF INVENTION: HIGH AFFINITY TGFA NUCLEIC
TITLE OF INVENTION: ACID LIGANDS AND INHIBITORS
NUMBER OF SEQUENCES: 89
CORRESPONDENCE ADDRESS:
ADDRESS: Swanson and Bratschun, L.L.C.
STREET: 8400 East Prentice Avenue, Suite #200
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,424B
FILING DATE: 2-JUNE-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/714,131
FILING DATE: 10-JUNE-1991
PRIOR APPLICATION DATA: 07/536,428
APPLICATION NUMBER: 07/536,428
```

FILING DATE: 11-JUNE-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/964,624
FILING DATE: 21-OCTOBER-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/117,991
FILING DATE: 8-SEPTEMBER-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/931,473
FILING DATE: 17-AUGUST-1992
ATTORNEY/AGENT INFORMATION:
NAME: Diane H. McLearn
REGISTRATION NUMBER: 33,960
REFERENCE/DOCKET NUMBER: NEX 34-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-458-424B-78

Query Match
Best Local Similarity 0.5%; Score 24.6; DB 1; Length 71;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgtgattagctgtattcaagcagatcaccattgcgatacgtcttc 4012
Db 56 GAAGACGTAGTAGTAGTAGTATTAACATACCCCCCTGGCCGCGATGCTCTCC 2

RESULT 7
US-08-973-124-78/c
Sequence 78, Application US/08973124
Patent No. 6207816
GENERAL INFORMATION:
APPLICANT: LARRY GOLD et al.
TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
TITLE OF INVENTION: LIGANDS TO GROWTH
NUMBER OF INVENTIONS: FACTORS
NUMBER OF SEQUENCES: 304
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,124
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08014
FILING DATE: 30-MAY-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,423
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,424
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,594

FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,591
FILING DATE: 05-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,725
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/479,783
FILING DATE: 07-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/518,693
FILING DATE: 20-MARCH-1996
ATTORNEY/AGENT INFORMATION:
NAME: Barry J. Swanson
REGISTRATION NUMBER: 33,215
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 793-3333
TELEFAX: (303) 793-3433
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 71 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-973-124-78

Query Match
Best Local Similarity 0.5%; Score 24.6; DB 4; Length 71;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgtgattagctgtattcaagcagatcaccattgcgatacgtcttc 4012
Db 56 GAAGACGTAGTAGTAGTAGTATTAACATACCCCCCTGGCCGCGATGCTCTCC 2

RESULT 8
PCT-US96-08014-78/c
Sequence 78, Application PC/TUS9608014
GENERAL INFORMATION:
APPLICANT: LARRY GOLD; NEROLISA JANIC; STEVEN RINGQUIST; NIKOS
TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
TITLE OF INVENTION: GROWTH FACTOR (TGF), PLATELET-DERIVED
TITLE OF INVENTION: KEROTINOCYTE GROWTH FACTOR (hKGF)
NUMBER OF SEQUENCES: 304
CORRESPONDENCE ADDRESS:
ADDRESSEE: Swanson & Bratschun, L.L.C.
STREET: 8400 E. Prentice Avenue, Suite 200
CITY: Englewood
STATE: Colorado
COUNTRY: USA
ZIP: 80111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: WordPerfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08014
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,423
FILING DATE: 02-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/458,424
FILING DATE: 02-JUNE-1995

```

10 RESULT
11 US-08-400-256-9/C
12 ; Sequence 9, Application US/08400256
13 ; Patent No. 5750497
14
15 GENERAL INFORMATION:
16 APPLICANT: Havelund, Svend
17 APPLICANT: Halstrom, John
18 APPLICANT: Jonassen, Ib
19 APPLICANT: Andersen, Asger Sloth
20 APPLICANT: Markussen, Jan
21 TITLE OF INVENTION: ACYLATED INSULIN
22 NUMBER OF SEQUENCES: 49
23
24 CORRESPONDENCE ADDRESS:
25 ADDRESSEE: No. 57504970 No. 5750497disk of No. 5750497th America, Inc
26 STREET: 405 Lexington Avenue, 64th Floor
27 CITY: New York
28 STATE: New York
29 COUNTRY: United States of America
30 ZIP: 10174-6401
31
32 COMPUTER READABLE FORM:
33 MEDIUM TYPE: Floppy disk
34 COMPUTER: IBM PC compatible
35 OPERATING SYSTEM: PC-DOS/MS-DOS
36 SOFTWARE: Patentin Release #1.0, Version #1.25
37
38 CURRENT APPLICATION DATA:
39 APPLICATION NUMBER: US/08/400,256
40 FILING DATE: 03-MAR-1995
41 CLASSIFICATION: 514
42
43 ATTORNEY/AGENT INFORMATION:
44 NAME: Tarnblis, Elias J.
45 RESISTRATION NUMBER: 33,728
46 REFERENCE/DOCKET NUMBER: 3985,220-US
47 TELECOMMUNICATION INFORMATION:
48 TELEPHONE: 212-867-0123
49 TELEFAX: 212-878-9655
50
51 INFORMATION FOR SEQ. ID NO.: 9:
52 SEQUENCE CHARACTERISTICS:
53 LENGTH: 1100 base pairs
54 TYPE: nucleic acid
55 STRANDEDNESS: single
56 TOPOLOGY: linear
57
58 MOLECULE TYPE: DNA
59
60 US-08-400-256-9

```


[illegible]

RESULT 15
US-08-672-158A-8
; Sequence 8, Application US/08672158A
; Patent No. 5770371
; GENERAL INFORMATION:
; APPLICANT: Sheryl Thompson
; TITLE OF INVENTION: Modification of Cryptic Splice Sites In
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 57703710 No. 5770371disk of No. 5770371th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/672,158A
; FILING DATE: 27-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 4855,000-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 83 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-672-158A-8

Query Match 0.5%; Score 23.8; DB 1; Length 83;
Best Local Similarity 72.1%; Pred. No. 2.3e+03;
Matches 31; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1352 gtacacatgtctgactctattatagaccattgtgacaa 1394
||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 gtgacgtatgtgagttatagtctacatctgtgtgcca 59

Search completed: May 29, 2002, 12:56:56
Job time: 9665 sec

Wed May 29 16:25:21 2002

us-09-676-436-3.szlml100.rst

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 09:13:38 ; Search time 3518.59 Seconds
(without alignments)
19141.135 Million cell updates/sec

Title: US-09-676-436-3
Percent score: 4990
Sequence: 1 ctgaagactctccgatagta.....ggtttatttaggaagctc 4990

Scoring table: IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 297742

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estha:*
2: em_esthm:*
3: em_esthm:*
4: em_esthm:*
5: em_esthm:*
6: em_esthm:*
7: em_esthm:*
8: em_esthm:*
9: gb_estl:*
10: gb_estl:*
11: gb_estl:*
12: gb_estl:*
13: em_gss_hum:*
14: em_gss_hum:*
15: em_gss_hum:*
16: em_gss_vitl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	39.8	0.8	95	9	AA990067	AA990067 ua58f02.r
2	29.8	0.6	70	9	AA490364	AA490364 aa44c10.r
3	28.6	0.6	71	9	A1767078	A1767078 w192907.x
4	27.8	0.6	83	10	B0063432	B0063432 B0063432
5	27.8	0.6	83	9	A1360096	A1360096 qy89d09.x
6	27.8	0.6	94	10	B1322285	B1322285 kx19a08.y
7	27.4	0.5	100	9	AA154655	AA154655 mq61c08.x
8	27.4	0.5	98	10	A1054386	A1054386 q176e03.x
9	27.4	0.5	100	10	C53310	C53310 C53310 yu11
10	26.8	0.5	81	10	BE573242	BE573242 601333139
11	26.6	0.5	82	10	BF101686	BF101686 601753418
12	26.6	0.5	87	9	AA267457	AA267457 m290a03.x
13	26.4	0.5	95	9	AA946634	AA946634 cq38f11.s
14	26.4	0.5	95	10	R40354	R40354 EST000038.S
15	26.2	0.5	78	10	R40354	R40354 yf81d05.s1
16	26.2	0.5	93	9	AV551027	AV551027 AV551027
17	26.2	0.5	93	9	AV551027	AV551027 AV551027

18	26.2	0.5	93	10	T62011	T62011 yb97b12.r1
19	26	0.5	78	10	BI858071	BI858071 603384887
20	26	0.5	85	12	BI127076	BI127076 G-10g24.f
21	26	0.5	88	9	AA865746	AA865746 oh41n11.s
22	26	0.5	98	9	A1251245	A1251245 qv55412.x
23	25.8	0.5	78	9	A1035450	A1035450 ub47b05.x
24	25.8	0.5	83	9	AA915897	AA915897 oh86b08.s
25	25.8	0.5	91	9	A1340694	A1340694 lb33c05.x
26	25.6	0.5	91	9	A1894187	A1894187 mc67e06.x
27	25.6	0.5	73	9	A1142956	A1142956 oz58d01.x
28	25.6	0.5	85	9	A1086378	A1086378 oz44c01.x
29	25.6	0.5	86	9	AV532477	AV532477 AV532477
30	25.4	0.5	76	9	AA149995	AA149995 x032c06.x
31	25.4	0.5	89	12	A2961698	A2961698 2M0230B17
32	25.4	0.5	91	12	D19910	D19910 HUMCS00867
33	25.4	0.5	92	12	AF149669	AF149669 AC20a02.s
34	25.4	0.5	93	9	AA669624	AA669624 AC20a02.s
35	25.4	0.5	94	9	AA662095	AA662095 AC20a02.s
36	25.4	0.5	94	9	AA576542	AA576542 hm66c06.s
37	25.4	0.5	95	10	BS362479	BS362479 qd73a06.y
38	25.4	0.5	99	10	BE66720	BE66720 SMOVAFCR
39	25.2	0.5	64	12	A2574215	A2574215 326v808
40	25.2	0.5	81	9	AA451873	AA451873 zx16e01.s
41	25.2	0.5	88	9	A1339161	A1339161 q14a01.s
42	25.2	0.5	88	9	A1339161	A1339161 q14a01.s
43	25.2	0.5	88	9	C01741	C01741 HDMS000363
44	25.2	0.5	94	10	BI943038	BI943038 sn03b04.y
45	25.2	0.5	97	12	TA330B01P	TA330B01P T. Bruce1

ALIGNMENTS

RESULT 1
AA990067 95 bp mRNA linear EST 02-JUN-1998
ua58f02.r1 Scores: thymus_2nbmt Mus musculus CDNA clone
IMAGE11361691 5', mRNA sequence.
AA990067
AA990067.1 GI:3175431

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 95)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisler, S., Kucaba, T., Lacy, M., Le, M., Martin, U., Morris, M.,
Schellenberg, K., Stepien, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, J., Lennon, G., Soares, B., Willson, R. and
Waterston, R.

TITLE
JOURNAL
COMMENT
The WashU-HM1 Mouse EST Project
Unpublished (1996)
Contact: Maria M. Mouse EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES
source
/organism="Mus musculus"
/strain="C57BL/6J1090"
/db_xref="taxon:10090"
/clone="IMAGE:1361691"
/clone_id="Soares-Thymus_2NBMT"
/sex="male"

High quality sequence stop: 82.
Seq primer: -28m13 rev2 ET from Amersham
MGI:699731
Location/Qualifiers
1..95

Query Match	Best Local Similarity	0.88;	Score 39.8;	DB 9;	Length 95;
Matches	53;	Conservative	0;	Mismatches	22;
				Indels	0;
				Gaps	0;
Qy	4471	gaatgaagccctagtagaataatgacttggaaattccttaac	roactactgtatgtaata	4530	
Db	16	GAACCGAGCCCGTGGCTAGAGTGTGACACGATCCCGT	ACTACTGATGTAATA	75	
Qy	4531	tttacaataagactg	4545		
Db	76	TTTACATAAAGACTG	90		

[illegible]

Accession	Version	Sequence
AA490364		sequence.
AA490364.1	GI:2219537	

SOURCE human.

REFERENCE	AUTHORS	HILLIERT	11
1 (bases 1 to 70)	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		

TITLE	JOURNAL	COMMENT
1. Waterston, R. and Wilson, R. Washu-Merck EST Project 1997 Unpublished (1997)		Waterston, R., White, Y., Wyllie, B., and Wilson, R. 1997

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMACE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity
Seq primer: -28bm13 rev2 ET from Amersham
High quality genome

Source	Location/Qualifiers	Sequence stop:
FEATURES		

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:823794"
/clone_id="Soares_NHMHU.S1"
/clone_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pUT3D-Pac
(Pharmacia) with a modified polylinker; Site.1: Not I;

```

Query Match		0.6%	Score 29;	DB 9;	Length 70;
Best Local Similarity		71.7%;	Pred. No. 1.4e+04;		
Matches	38;	Conservative	0;	Mismatches 15;	Indels 0;
Gaps	0				
Qy	4051	aarattcttccctaacctcgcgtgataatcaacagtggagatttgatgttc	4103		
Dd	8	AATATTCTTCCTCACTCCAGAAATGAAGAATTTGGACACTTGTGATGC	60		

RESULT	3
LOCUS	A1767078
DEFINITION	A1767078 71 bp mRNA
VERSION	w192507.x1 NCI-CCAP-K1d12 Homo sapiens cDNA clone IMAGE:2400828 3'
ACCESSION	A1767078
KEYWORDS	EST.
SOURCE	bio.
GI:	5233587

ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Human.					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Euteria; Primates; Carnivora; Catarrhini; Hominoidea; Homo.					
1 (bases 1 to 71)					
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .					
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),					
Tumor Gene Index					
Unpublished (1997)					
Contact: Bob					

Email: robert.strausberg, Ph.D.
 Tissue Procurement: Christopher Moskalk, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Sequencing by: Greg Lennon, Ph.D.
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.dio.llnl.gov/bdbp/image/image.html
 Insert Length: 785 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence step: 70.
 Location/Qualifiers
 1
 71

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2400828"
/clone_lib="NCI_CGAP_Kid13"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
/notes="Organ: kidney; Vector: pUC19"

```

BASE COUNT 22 a 16 c 7 g 25 t

Plasmid DNA from the site-directed library NCI/COAP Kid5 was prepared, and ss circles were made *in vitro*. Following HAV hybridization, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs (clonoids 1323912-1325831, 1471568-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo.

us-09-676-436-3.szlm100.rst

Query Match	Score	DB %	Length	0;
Best Local Similarity	64.28;	0;	Mismatches 24;	Indels 0;
Matches 43;	Conservative			
Qy 3085	ttgagattctaaagcaattgaacccgtctatttcagctttacacgaagttacgtc	3144		
Db 2	tttttttgcgcacacacagatgaacttatttaattatcaaaatgcacgtgaaaaatgccttc	61		
Qy 3145	ttgagtt 3151			
Db 62	tttcattt 68			

[illegible]

FEATURES	SOURCE	BASE COUNT	ORIGIN
17	Loc843	15	a
17	organism="Xenopus laevis"	11	c
17	/db_xref="taxon:8355"	11	g
17	/clone="XL076e13"	42	t
17	/clone_lib="NIBB Kocchi normalized Xenopus tailbud library"	4	others
17	/tissue_type="whole embryo"		
17	/dev_stage="stage 25"		

Query Match	0.68;	Score 28;	DB 10;	Length 83;	
Best Local Similarity	63.38;	FreeE No. 2.6e+04;			
Matches	40;	Conservative	0;	Mismatches 23;	Indels 0;
135	tttgaagaagaaatgatgtgacccaatcagctccacataaagaacatggaaaacagt	194			
75	tggnnaaaaaaaaaatgcaacnnaaaaaaaaaaacatccatttttaaaacatttctcacacaact	16			
195	gga	197			
15	gga	13			
Db					
RESULT - 5					
AI360096/c	AI360096	83 bp	mrna	linear	EST 16-FEB-1996
LOC4989d09.x1	NCI_CGAP_Brn25	Homo sapiens	cdna	clone	IMAGE:2019185 3'
DEFINITION	mrna sequence.				
AI360096	AI360096				
VERSION	AI360096.1	GI:4111717			
KEYWORDS	EST.				

SOURCE	Mammals
ORGANISM	Chordata; Craniata; Vertebrata; Euteleostomi; Euryarchontes; Metazoa; Primates; Carnivora; Homnidae; Homo
REFERENCE	1 (bases 1 to 83) http://www.ncbi.nlm.nih.gov/ncicgap
AUTHORS	National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (NCICGAP), Tumor Gene Index
JOURNAL	Unpublished (1998)
COMMENT	Contact: Robert Stransberg, Ph.D. Email: cagabb@ismail.nhlbi.gov Louis, M.D., Myrna R. Rosenfeld M.D.

Bernaldo, Ph.D.
 Sequencing Center
 DNA Sequencing
 Restriction: NCI-CGAP clone distribution: NCIMBL/BLN at:
 found through the I.M.A.G.E. Consortium/BLN at:
 www-bio.1nl.gov/Db/rf/imag/blast/blast.html
 Insert Length: 558 bp/imag/blast/blast.html
 Seq Primer: 159 bp/imag/blast/blast.html
 High qual location/Qualifiers

```

FEATURES
SOURCE
1..83
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1="IMAGE:219185"
/clone_1p="M-219185.CGAP.Brn25"
/clone_1bp="M-anaplastic oligodendroglioma"
/tissue_type="PB108"
/lab_host="organ: brain; Vector: pTV73D-Pac (Pharmacia) with a
/notestrain; Site:1: Not I; Site:2 Eco RI; 1st
/modified polylinker; Site:1: Not I; Site:2 Eco RI; primer 15',
stranded cDNA was primed with a Not I and Eco RI
13311" double-stranded cDNA was ligated to Eco RI
13311" double-stranded cDNA was ligated to Eco RI
adaptors (pharmacia), digested the modified pTV73 vector.
the Not I and Eco RI sites of was constructed by Bento
library is normalized, and was constructed by Bento
library and M.Patima Bonaldo.
19 a 6 c 4 g 54 t

```

	Query Match	0.6%	Score 27.8;	DB 9;	Length 83;
	Best Local Similarity	59.5%	Fred. No 3e+04;	Indels 0;	Gaps 0;
	Matches 47; Conservative	0;	Mismatches	32;	
ly	475 ttggaactctacgacttgcctaagcttcctccacgtccaagaanaaacaggag	554			
	83 TTTAAAAATTTTAAATTTCTCCTTAAGTCACGCTGGGTGTTAAGAAAAAAAAAAAAA	24			
dDb	535 caaagagagacaagaata	553			
dDb	23 AAAAAAAAAAAAAAAAANA	5			
RESULT 6					
B1322285	LOCUS	94 bp	mRNA	linear	EST 30-Jul-2001
DEFINITION	K132608 v3 Parastromyloides trichosuri FL pMMP1 v1 Chiapelli				
ACCSSION	B1322285				
VERSION	B1322285.1				
KEYWORDS	EST.				
SOURCE	Parastromyloides trichosuri.				
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhadditida;				
	Panagrolaimoidea; Stromyloididae; Parastromyloides.				
REFERENCE	1 (bases 1 to 94)				
AUTHORS	McCartter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marre,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y				


```

/clone_image=11772
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT73d (Pharmacia)
with a modified polylinker; Site_1: Pac I; site_2: Eco RI;
1st strand cdna was primed with a Pac I - oligo(dN) primer
15' AATCGGAGACGAATTAAATTAAGAATCTTTTCTTTTTTTTTTTT 3',
double-stranded cdna was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Patima Bonaldo."

```

RESULT	9
LOCUS	A1054386/c
DEFINITION	g176e33x.t1 NCL_CGAP_Ov26 Homo sapiens cDNA clone IMAGE1862428 3' , mRNA linear EST 27-JUL-1996 98 bp
ACCESSION	A1054386
VERSION	A1054386.1
KEYWORDS	EST
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap/ .
TITLE	(Bases 1 to 98)
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor gene Index
COMMENT	Unpublished (1997) Contact: Robert Strausberg, Ph.D. Email: cgapsb@remail.nhi.gov Tissue Procurement: Melissa Brown, M.D., Elise Kohn, M.D., Michael R. Emmert-Buck, M.D., Ph.D.

```

FEATURES
source
1. .98
Location/Vacuole
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image:1862428"
/clone_11b="NCL_CGAP_OV26"
/sex="female"
/tissue_type="papillary serous carcinoma"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: ovary; Vector: pAPM1; mRNA made from
papillary serous ovarian carcinoma, cdna made by oli90-dt
priming. Directionally cloned. Size-selected on agarose
gel, average insert size 500 bp. Primary library,
non-amplified."
BASE COUNT
6 a 28 c 7 g 57 t
ORIGIN

```

```

Query Match      0.5%; Score 27.4; DB 9; Length 98;
Best Local Similarity 62.3%; Pred. No. 4.1e+04; Indels 0; Gaps 0;
Matches 43; Conservative 0; Mismatches 26;

QY 1127 aggttgatgacacagaggaattaaagaggttcgaagtgtcagatgagtgagtgag 1186
DB 75 AGCGGGAGGAAGAAATCGAAAGAGATTAGAGAGGACACCAAGATTAATACGAGAGCAAAATA 16
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1187 aagagacaa 1195
DB 15 AAAAAAAAA 7
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10      100 bp      mRNA      linear      EST 11-SEP-1997
LOCUS C53310
DEFINITION C53310 yuji Kohara unpublished cDNA Caenorhabditis elegans CDNA
ACCESSION C53310
VERSION C53310
KEYWORDS clone yk259b3 3', mRNA sequence.
SOURCE C53310.1 GI:2391067
ORGANISM Est.
          Caenorhabditis elegans.
          Caenorhabditis elegans.
          Nematoda; Chromadorea; Rhabditida; Rhabditoidea.
          Bursaroya; Metazoa; Pelodermata; Caenorhabditis.
          1 (bases 1 to 100)
          Kohara,Y., Mochoshiri,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
            M., Miyata,A. and Nishigaki,A.
            Expression map of the C.elegans genome
            Unpublished (1996)
            Contact: Yuji Kohara
            Genome Biology Lab.
            National Institute of Genetics
            Yata 111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
            Location/Qualifiers
              location=1..100
              organism="Caenorhabditis elegans"
              strain="Cb1489 him-8(e1489)"
              db_xref="taxon:6239"
              clone_yk259b3"
              clone_lib="yuji Kohara unpublished cDNA"
              sex="hermaphrodite, male"
              tissue_type="whole animal"

```


Wed May 29 16:25:21 2002

us-09-676-436-3.sz1m100.rst

Page 8

Email: M-Saber@RCU.EUN.BG
Seq primer: sk

FEATURES

location/Qualifiers

1..95
/organism="Schistosoma mansoni"
/strain="Egyptian"
/db_xref="taxon:6183"
/clone="SWTADAMS0038SK"
/clone_lib="S. mansoni cDNA"
/lab_host="E. coli XL Blue1"
/note="vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; mRNA was purified from adult couples of S. mansoni. vector priming with the pBluescript II SK+ vector. cDNA was directionally synthesized from the EcoRI site in the vector to the XhoI site." 31 a 15 c 10 g 39 t

BASE COUNT

31 a 15 c 10 g 39 t

Query Match

Best Local Similarity 59.2%; Score 26.4; DB 10; Length 95;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 641 actctcttatacaagccgtaagccatccagatatataatgaatcctactt 700
||| || |||| | || |||| ||| |||| |||| |||| |||| ||||
Db 5 ACTATGATTATATACACTACTTCAGCAATCCCTATCTATAAATGTATCTTATTT 64
QY 701 tcaagtcgactatgg 716
| ||| | ||| |
Db 65 TGAATAATTTACCATAG 80

Search completed: May 29, 2002, 11:17:02
Job time: 7404 sec

Wed May 29 16:25:19 2002

us-09-676-436-3.sz1m100.ing

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 10:21:32 ; Search time 487.5 Seconds
without alignments
17574.160 Million cell updates/sec

Title: US-09-676-436-3
Perfect score: 4990
Sequence: 1 ctgaagactctccgatga.....gtttatttagaagctc 4990

Scoring table: IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 2046006

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N.Geneseq.032802.*
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001.DAT.*
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27.2	0.5	78	18 AAT50934	Mouse p53-recognition
2	26.2	0.5	92	21 AAC19445	Human secreted pro
3	25.2	0.5	100	16 AAC08400	Human insulin deri
4	25.2	0.5	60	22 AAH48689	Human G-protein su
5	25.2	0.5	70	21 AAA71594	Human brain natru
6	25.2	0.5	89	16 AAT22086	Human gene signatu
7	25.2	0.5	99	18 AAT92244	Mercuric ion reduc
8	25.2	0.5	93	20 AAZ28594	Primer 307-3395 fo
9	25	0.3	54	21 AAA73946	GFP Leu(CITG) 5 forw

10	25	0.5	91	22 ABA69691	Human foetal liver
11	24.8	0.5	77	21 AAC12392	Human secreted pro
12	24.8	0.5	92	21 AAS06930	Reverse PCR primer
13	24.8	0.5	94	21 AAC11499	Human secreted pro
14	24.8	0.5	94	21 AAC15541	Transforming growt
15	24.6	0.5	71	18 AAT65208	CDNA encoding pros
16	24.6	0.5	72	21 AAT52344	Human neutrophil c
17	24.6	0.5	98	20 AAX23469	Human foetal liver
18	24.4	0.5	82	20 AAB68515	Human foetal liver
19	24.4	0.5	99	22 ABA68515	Probe #14054 for g
20	24.4	0.5	99	22 ABA68515	Human insulin deri
21	24.4	0.5	100	16 AAO86811	Human foetal liver
22	24.2	0.5	80	22 ABA72501	Human brain expres
23	24.2	0.5	80	22 ABA72501	Human bone marrow
24	24.2	0.5	80	22 AAK47388	Human bone marrow
25	24.2	0.5	80	22 AAK47388	Probe #21912 used
26	24.2	0.5	84	21 AAA92397	Rhymosin alpha-1 n
27	24.2	0.5	94	21 AAA92397	Sbq1 exon sequence
28	24.2	0.5	94	21 AAA92397	Sbq1 exon sequence
29	24.2	0.5	94	21 AAA92397	Human foetal liver
30	24.2	0.5	95	22 ABA71915	Human brain expres
31	24.2	0.5	95	22 ABA71915	Human bone marrow
32	24.2	0.5	95	22 ABA71915	Probe #20956 used
33	24.2	0.5	95	22 ABA71915	Human foetal liver
34	24.2	0.5	95	22 ABA71915	Probe #18368 for g
35	24.2	0.5	99	22 ABA75243	Human foetal liver
36	24.2	0.5	99	22 ABA75243	Human brain expres
37	24.2	0.5	99	22 AAK49875	Human bone marrow
38	24.2	0.5	99	22 AAK49875	Human bone marrow
39	24.2	0.5	99	22 AAI26980	Probe #24484 used
40	24.2	0.5	66	22 AAD14229	Synthetic transcri
41	24.2	0.5	78	24 AAS63113	Cell death protect
42	24.2	0.5	97	14 AAO37278	Clone 92-2 used in
43	24.2	0.5	97	16 AAO94997	SP 5.11.11.5 clon
44	24.2	0.5	97	19 AAV35835	Synthetic storage
45	24.2	0.5	97	20 AAV95532	Synthetic lysine-r

ALIGNMENTS

RESULT 1
AAT50934 standard; CDNA: 78 BP.
ID AAT50934:
AC AAT50934:
XX 07-Apr-1997 (first entry)
XX Mouse p53-recognition clone 1 5' sequence.
XX p53 binding protein; Mdmx; tumour suppressor; cancer; ss.
XX Mus sp.
XX WO9641875-A1.
XX 27-DEC-1996.
XX 13-JUN-1996; 96WO-NL00239.
XX 13-JUN-1995; 95EP-0201565.
XX (INTR-) INTRIGENE BV.
XX (OYLE-) RIKSUNIV LEIDEN.
XX Jochimsen A, Shwarts A, Van Der Eb AJ;
XX WPI; 1997-065462/06.
XX Nucleic acid encoding p53-binding protein - which has homology to
XX mouse Mdm2, for use in cancer research

PS Disclosure; Fig 2a; 30pp; English.
XX

CC were dert. for cDNA clone 1, isolated from a mouse 16-day embryo
CC cDNA library using tumour suppressor p53 as probe. The 5' and 3'
CC regions (AA150935-37) of another clone were also dert. These
CC sequences showed no significant homology to any known sequence
CC submitted to several databass. A third clone (see also AA150933)
CC coded for a p53 binding protein designated Mdmx (AA110206).
XX
SQ Sequence 78 BP; 33 A; 9 C; 22 G; 10 T; 4 other:

Query Match	0.5%	Score 27.2;	DB 18;	Length 78.8;
Best Local Similarity	58.1%;	Pred. No. 1e+03;		
Matches 43;	Conservative 2;	Mismatches 20;		

[illegible]

RESULT	2
AAC19445/c	
ID	AAC19445 standard; cDNA; 92 BP
XX	

06-UCJ-2000 (first entry)
D1
XX

Human secreted protein 5' EST, SEQ ID NO: 23520.

gene therapy; chromosome mapping; ss. CDNA isolation;

Homo sapiens.

EP1.033401-A2.

06-SEP-2000.

41-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

X - James Milne Edwards J, Duclert A, Giordano J;

WFL; 2000-500381/45.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures.

Claim 1; SEQ ID 23520; 71pp + CD-ROM; English.

The present sequence is one of a large number of 5' ESTs derived from mRNAs encoding secreted proteins. No ORF has yet been conclusively identified within the present sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs derived from 30 different tissues. EST sequences usually correspond mainly to the 3' untranslated region (UTR) of the mRNA because they are often obtained from oligo-dT primed cDNA libraries. Such ESTs are not well suited for isolating cDNA sequences derived from the 5' ends of mRNAs and even in those cases where longer cDNA sequences have been obtained, the full 5' UTR is rarely included. 5' ESTs are derived from RNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures.

CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX

sequence 34 BF; 12 A; 9 C; 15 G; 47 T; 9 other;

Query Match	0.5%;	Score 26.6;	DB 21;	Length 92;
Best Local Similarity	57.1%;	Pred. No. 1.7e+03;		
Matches	36;	Conservative	6;	Mismatches 31;

[illegible]

RESULT	3
AAQ86400/c	
ID	AAQ86400 standard; DNA; 100 BP.
XX	

16-NOV-1995 (first entry)
D1
XX

XX derivative AluA1 A-chain PCR primer #128

ala21 PCR primer #128; ss.

Synthetic.

W09507931-A.

23-MAR-1995.

16-SEP-1994; 94WO-DK00347.

17-SEP-1993; 93DK-0001044.
02-FEB-1994; 04TC 0100000

02-FEB-1994; 94US-0190829.

(NOVO) NOVO-NORDISK AS

Markussen J; Havelund S, Jonassen I;

WPI; 1995-131314/17.

complex - is used to treat diabetes and is rapid acting.

Example 1; Page 30; 100pp; English.

CC The Asp3-B-chain/A1a21 A-chain human insulin derivative DNA.
CC The two subunits of the derivative are covalently connected via
CC disulphide bonds between Cys residues A7/B7 and A20/B19. The
CC derivative, which may be present as a zinc ion complex, can
CC be used as a fast action treatment for diabetes.
XX

sequence 100 Bp; 29 A; 22 C; 23 G; 26 T; 0 other,

Query Match	0.5%;	Score 26;	DB 16;	Length 100;
Best Local Similarity	55.6%;	Pred. No. 2.7e+03;		
Matches 50:	Conservative	0;	Mismatches 40;	

[illegible]


```

PF 21-APR-1995; 95US-0427097.
XX
XX
PR 21-APR-1995; 95US-0427097.
XX
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
XX
PI Meagher RB, Summers AO;
XX
XX
DR WPI; 1997-470112/43.
XX
XX
PT DNA encoding mercuric ion reductase in plant-expressible form -- for
PT producing transgenic plants resistant to toxic heavy metals
XX
XX
PS Example 1; Column 43-44; 52bp; English.
XX
XX
CC The present sequence represents a PCR mutagenesis primer used to
CC mutagenise the mera coding sequence derived from transposon Tn21, to
CC adapt it for plant expressibility. The nucleic acid molecule produced
CC is useful for producing transgenic plants that are resistant to toxic
CC heavy metals (especially mercury) and so can be used to remediate and/or
CC revegetate contaminated areas. The bacterial mera gene is derived from
CC the transposon Tn21, which was originally isolated from the
CC incompatibility group Incrit resistance plasmid NR1.
XX
XX
SQ Sequence 99 BP; 29 A; 21 C; 25 G; 24 T; 0 other;

Query Match 0.5%; Score 25.2; DB 18; Length 99;
Best Local Similarity 54.3%; Pred. No. 4.7e+03;
Matches 51; Conservative 0; Mismatches 43; Indels 0; Gaps

QY 3026 atgtctgactcaatgtgagagtgtagaggtacaagaccagctggcgactcaagat 3085
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 aagaccagcatatggtggaagctgtactgcgtcattcgactgtgaagcattgaagtc 61
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 3086 ttgatcttccaagcattgacccgccttat 3119
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 gtgagcatactcaagcaagcgaattgctctatat 95

RESULT 8
ID: AA228594 standard; DNA; 99 BP.
XX
XX
AC AA228594;
XX
XX
DT 24-DEC-1999 (first entry)
XX
XX
DE Primer 307-339S for Tn21 mera gene.
XX
XX
KW Organometal; resistance; transcription; translation; regulation; mera;
KW transposon; Tn21; merB; mercuric ion reductase; organomercury lyase; ss;
KW transgenic plant; detoxification; metal compound; soil; sediment; primer;
KW aquatic environment; bioremediation; contamination; PCR; amplification.
XX
XX
OS Synthetic.
OS Transposon Tn21.
XX
XX
PN US5965796-A.
XX
XX
PD 12-OCT-1999.
XX
XX
PF 19-JUN-1997; 97US-0878957.
XX
XX
PR 21-APR-1995; 95US-0427097.
XX
XX
PA (UYGE-) UNIV GEORGIA RES FOUND INC.
XX
XX
PI Rough CL, Meagher RB, Summers AO;
XX
XX
DR WPI; 1999-579950/49.
XX
XX
TT Metal resistance sequences for producing transgenic plants -

```

Example 3; Column 17-18; 71pp; English.

The invention relates to a nucleic acid molecule comprising a coding sequence for an organometal resistance protein, which is operably linked downstream of and under the regulatory control of a plant-expressible transcription and translation regulatory sequence. The organometal resistance gene is especially the merA gene from the transposon Tn21 or the E.coli merB gene. merA encodes a mercaptic ion reductase and merB encodes a organomercury lyase. The nucleic acid sequences are useful for producing transgenic plants which are capable of efficiently processing mercury. These plants are able to remove or detoxify metal compounds such as methyl mercury and ionic mercury from soil, sediment, and aquatic environments contaminated sites. Bioremediation of metal and organometal contaminated sites. This sequence represents a PCR primer used to amplify the Tn21 merA gene which is mutated by random mutagenesis (see AA228584-228589 for mutated sequences) and the mutated sequences are expressed in the transgenic plants of the invention.

Sequence 99 BP; 29 A; 21 C; 25 G; 24 T; 0 other;

	Query Match	Best Local Similarity	Score	Pred.	No. Mismatches	Indels	Gaps
OY	3026 atctcgtgactgaatgtagagtgtgtagagtcacagaccagggtggcgactcaaggat	54.3%	25.2	4.7e+03	43	0	0
DB	2 aagaccagcatagaagtgcgttgcattcgcattccagcaggaagcattgaaatgc	51; Conservative	0	0	0	0	0
OY	3086 ttgattttccaagaacattgaacctgttat	54.3%	25.2	4.7e+03	43	0	0
DB	62 gtgcgcatactacaagcaagcgaagtgcctatat	51; Conservative	0	0	0	0	0

RESULT 9
AAAT3946/C
ID AAA73946 standard; DNN: 54 BP.
AC AAA73946;
DX 06-DEC-2000 (first entry)
DE GFP Lcu(CTG)5 forward primer.
KW Green fluorescent protein; GFP; reporter gene; codon utilisation;
KM translational efficiency; protein abundance; PCR primer; ss.
XX Synthetic.
XX WO200042215-A1.
PN 20-JUL-2000.
PD 07-JAN-2000; 2000WO-AU00008.
PF 08-JAN-1999; 99AU-0008078.
KA (UYOU) UNIV QUEENSLAND.
PA (SUNK/) SUN X.
PR Zhou J, Frazer IH;
PI WPI: 2000-499118/44.
DR
XX Determining translational efficiency of codons in cells, comprising
XX introducing synthetic constructs with reporter genes fused in frame to
XX tandem repeats of the codon, and measuring expression -
PS Example 1; Page 183; 190pp; English.
XX The present sequence is a primer used to generate a synthetic gfp gene by

CC PCR amplification of a humanised gfp gene. A single artificial start
CC codon followed by a stretch of five identical codons was fused in frame
CC immediately upstream of a gfp coding sequence to form the synthetic gene.
CC The amplified fragment was cloned into the mammalian expression vector
CC pCDNA3, which contains SV40 ori and the CMV promoter, and was used in a
CC method for determining the translational efficiency of a codon in a cell.
CC The synthetic construct was introduced into COS-1 cells and expression of
CC the reporter protein (green fluorescent protein) was measured. A series
CC of 64 gfp reporter constructs was made in which the gfp gene is preceded
CC in frame by a tandem string of 5 identical codons. The series covers the
CC entire set of isotocpept of 5 identical codons. Codons with a higher
CC translational efficiency than their corresponding synonymous codons can
CC be identified. These codons may then be used to replace the less
CC preferred codons of a polynucleotide so that there is higher protein
CC expression within undifferentiated epithelial cells such as COS-1 cells.

CC Sequence 54 BP; 9 A; 14 C; 20 G; 11 T; 0 other;

XX

SO

Query Match 0.5%; Score 25; DB 21; Length 54;
Best Local Similarity 69.4%; Pred. No. 3.6e+03;
Matches 34; Conservative 0; Mismatches 15; Indels 0; Gaps 0

OY 2732 ccagtcagccggatcatgcccaaacgtttgcacacacgtaagaatgaatgc 2780
||||| + | | | | | | | | | | | | | | | | | | |
Db 53 CCAGTGAACACTGCTCCTGCCCTTGTCACACACGACGACGATCATGTAC 5

RESULT 10
ABA69691/C
ID ABA69691 standard; DNA; 91 BP.
XX
AC ABA69691;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #17996.
XX
KA Human; foetal liver; gene expression; single exon nucleic acid probe; ss
XX OS Homo sapiens.
XX FN NC0200157277-A2.
XX PD 09-AUG-2001.
XX PE 30-JAN-2001; 2001WO-US00669.
XX PF 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 23-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WP: 2001-483447/52.
XX DR WPI: 2001-483447/52.
XX PS Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human fetal liver -
XX Claim 4: SEQ ID NO 17996; 639bp + sequence listing; English.
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX foetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.

CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SO Sequence 91 BP; 28 A; 19 C; 13 G; 31 T; 0 other;

Query Match 0.5%; Score 25; DB 22; Length 91;
 Best local Similarity 61.5%; Pred. No. 5.1e+03;
 Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 3815 acgaatgaaatcgaagcacaacaccccaatcgtgctgattgtgctg 3874
 Db 67 AAGAAATTTTGTGATGAAAGACTGTAAACACCAATATTGCTTATTGGAGCT 8

OY 3875 agctc 3879
 Db 7 ATCTC 3

RESULT 11
 AAC12392/c
 ID AAC12392 standard; cDNA; 77 BP.
 XX
 AC AAC12392;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 16467.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;
 DR WPI: 2000-500381/45.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 1; SEQ ID 16467; 71pp + CD-ROM; English.

CC The present sequence is one of a large number of 5' ESTs derived from
 CC cDNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are often obtained from oligo-dT primed cDNA
 CC libraries. Such ESTs are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.

SO Sequence 77 BP; 27 A; 14 C; 15 G; 21 T; 0 other;

Query Match 0.5%; Score 24.8; DB 21; Length 77;
 Best local Similarity 60.3%; Pred. No. 3.2e+03;
 Matches 41; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

OY 301 gactcagtggtgagcttcgtcccaaaatcaatccagatgagatcaatagct 360
 Db 74 GACTGTAGCTGCTCAATTTTAAAAATATTAACCTCCCATTTGCTCTGGATTAAGCT 15

OY 361 taccctcag 368
 Db 14 TATCTCAG 7

RESULT 12
 AAS06930/c
 ID AAS06930 standard; DNA; 92 BP.
 XX
 AC AAS06930;

DT 12-SEP-2001 (first entry)

DE Reverse PCR primer for RSV A G gene insert.

KW Infectious chimeric parainfluenza virus; antigenic determinant;
 KW nucleocapsid phosphoprotein; large polymerase; attenuated vaccine;
 KW human PIV1; HPIV1; HPIV2; RSV; pathogen; measles; PCR primer;
 KW respiratory syncytial virus; respiratory tract infection; bovine; ss.

OS Human respiratory syncytial virus.
 PN WO200142445-A2.

PD 14-JUN-2001.

PF 08-DEC-2000; 2000WO-US33293.

PR 10-DEC-1999; 99US-0170195.

PR 10-DEC-1999; 99US-0458813.

PR 10-DEC-1999; 99US-0459062.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Murphy BR, Collins PL, Schmidt AC, Durbin AP, Skiadopoulos MH;
 PI Tao T;

DR WPI: 2001-356173/37.

PT Isolated infectious chimeric parainfluenza virus (PIV), useful in an
 PT attenuated vaccine to elicits an immune response against one or more
 PT virus(es) selected from human PIV1 (HPIV1), HPIV2 and HPIV3 -
 PS Example 10; Page 150; 305pp; English.

CC The present sequence for reverse PCR primer for respiratory syncytial
 CC virus (RSV) A G gene insert is used with the forward PCR primer
 CC (AAS06929) in the construction of bovine/humanPIV3-RSV chimeric viruses.
 CC The sequence is described in an invention relating to novel infectious
 CC chimeric parainfluenza viruses (PIVs). The virus comprises a major
 CC nucleocapsid protein (N), a nucleocapsid phosphoprotein (P), a large
 CC polymerase protein (L), and a partial or complete PIV vector background
 CC genome, or antigenome combined with one or more heterologous gene(s) or
 CC more heterologous pathogen(s) to form a chimeric genome or antigenome.
 CC The chimeric PIV is useful in an attenuated vaccine to elicit an immune
 CC response against one or more virus(es) selected from human PIV1 (HPIV1),
 CC HPIV2 and HPIV3. The chimeric PIV may also elicit a polyspecific immune
 CC response against HPIV3, measles or RSV. An immunospecific composition
 CC may also contain two chimeric PIVs, where the first chimeric PIV elicits
 CC an immune response against HPIV3 and the second chimeric PIV elicits
 CC immune response against HPIV1 or HPIV2, and where both the first and
 CC second chimeric PIVs elicit an immune response against the non-PIV
 CC pathogen. Chimeric HPIV3, HPIV1 and HPIV2 are useful as vaccines to
 CC prevent measles and upper or lower respiratory tract infections


```

RESULT 15
ID AAT65208/c
XX AAT65208 standard; DNA: 71 BP.
AC AAT65208;
XX
DT 10-SEP-1997 (first entry)
XX
DE Transforming growth factor beta-1 binding ligand D 11.
XX
KW Transforming growth factor; beta-1; TGF-beta-1; binding ligand;
KW identification; SELEX; anti-mitogenic; inhibition; cell;
KW Systematic Evolution of Ligands by EXponential enrichment;
KW epithelial; proliferation; diagnosis; treatment; fibroids;
KW kidney; lung; liver; dermal scarring; restenosis; ss.
XX
OS Synthetic.
XX
PN WO9638579-A1.
XX
PD 05-DEC-1996.
XX
PF 30-MAY-1996; 96WO-US08014.
XX
PR 20-MAR-1996; 96US-0618693.
PR 02-JUN-1995; 95US-0458423.
PR 02-JUN-1995; 95US-0458424.
PR 05-JUN-1995; 95US-0465591.
PR 05-JUN-1995; 95US-0465594.
PR 07-JUN-1995; 95US-0479725.
PR 07-JUN-1995; 95US-0479783.
XX
PA (NEXS-) NEXSTAR PHARM INC.
XX
PI Gold L, Janjic N, Pagratlis N, Ringquist S, Toolman PJ;
XX WPI; 1997-034387/03.
XX
PT Identification of nucleic acid ligands to TGF-beta, PDGF and hKGF
PT using SELEX, used in the diagnosis and treatment of proliferative
XX disorders
XX
PS Claim 15; Page 123; 209pp; English.
XX
CC The present sequence, a transforming growth factor beta-1
CC (TGF-beta-1) binding ligand, was identified by Systematic Evolution
CC of Ligands by Exponential enrichment (SELEX). Briefly a candidate
CC mixture of nucleic acids was contacted with TGF-beta-1, and nucleic
CC acids having an increased affinity to TGF-beta-1 partitioned from
CC the remainder of the mixture. The partitioned nucleic acids were
CC then amplified to yield a mixture of nucleic acids enriched for
CC sequences with higher affinity and specifically for binding to
CC TGF-beta-1. The ligand is anti-mitogenic and may be used to inhibit
CC epithelial cell proliferation, or in the diagnosis and treatment of
CC TGF-beta-1 mediated pathological conditions, e.g. fibrotic
CC conditions such as fibroids of the kidney, lung and liver and more
CC acute conditions such as dermal scarring and restenosis.
XX
SO Sequence 71 BP: 15 A; 16 C; 22 G; 18 T; 0 other;

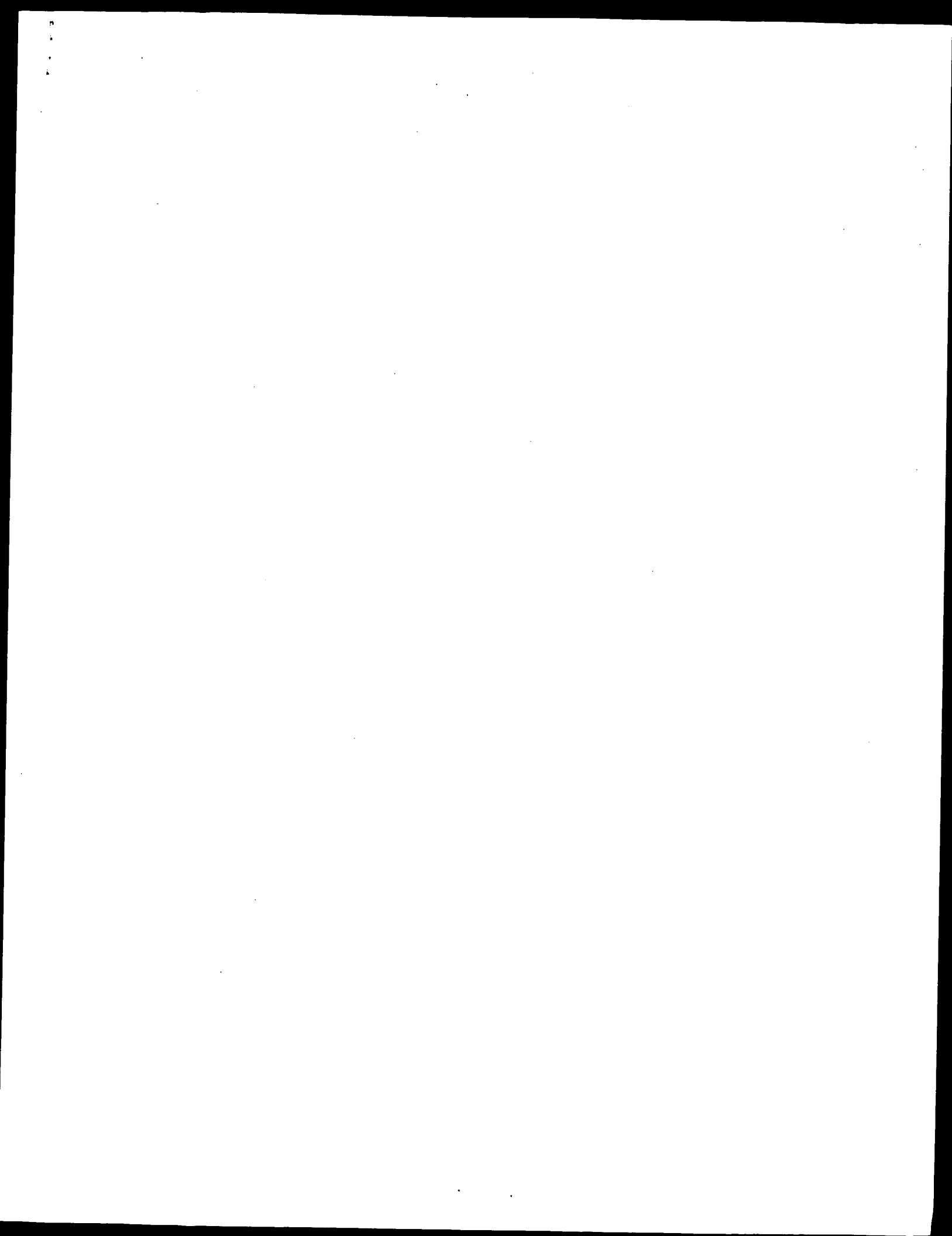
Query Match 0.5%; Score 24.6; DB 18; Length 71;
Best Local Similarity 65.5%; Pred. No. 5.7e+03;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 3958 gaacatgtattagctgattcaagaagcagatcaccattgcatacagctctcc 4012
DB 56 GAAGACGTAGTAAAGTAAATTAACATACCCCTGGCCGCGATCGTCCTCC 2

```


Wed May 29 16:25:19 2002

us-09-676-436-3.sz1ml100.rng



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 29, 2002, 09:42:16 ; Search time 5865.16 Seconds
(without alignments)
17804.019 Million cell updates/sec

Title: US-09-676-436-3
Sequence: 4990 1 ctgaagaactctcccgatga.....ggtttattttagaagctc 4990

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 843946

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: 1: gb_da: 2: gb_htg: 3: gb_in: 4: gb_cm: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_pr: 10: gb_ro: 11: gb_sy: 12: gb_un: 13: gb_vl: 14: gb_vl: 15: em_da: 16: em_fun: 17: em_hum: 18: em_in: 19: em_mu: 20: em_ov: 21: em_or: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vl: 30: em_htg_hum: 31: em_htg_inv: 32: em_htg_other: 33: em_htg_inv:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

C	1	29.6	0.6	87	3	DDIDDKD	M59747 D. discoideum
	2	26	0.5	100	6	AR007403	AR007403 Sequence
	3	25.4	0.5	96	5	AF025984	AF025984 Paratrich
	4	25.2	0.5	70	1	AX033167	AX033167 Sequence
	5	25.2	0.5	70	6	AX033159	AX033159 Sequence
	6	25.2	0.5	99	6	165769	165769 Sequence 5
	7	25	0.5	83	4	RABPFKM18	M14473 Rabbit musc
	8	24.8	0.5	92	6	AX173377	AX173377 Sequence
	9	24.6	0.5	71	6	AR140801	AR140801 Sequence
	10	24.6	0.5	71	6	193480	193480 Sequence 78
	11	24.6	0.5	71	6	193107	193107 Sequence 78
	12	24.4	0.5	73	3	S76509	S76509 D18 (A) fh
	13	24.4	0.5	98	6	DROTFRANSIN	120950 Drosophila
	14	24.4	0.5	100	3	AR007408	AR007408 Sequence
	15	24	0.5	51	10	U92173	U92173 Mus musculu
	16	24	0.5	66	6	AX207310	AX207310 Sequence
	17	24	0.5	88	9	AB04840511	AB048411 Homo sapi
	18	24	0.5	97	6	AR014621	AR014621 Sequence
	19	24	0.5	97	6	BD010413	BD010413 Chimeric
	20	24	0.5	97	6	I26754	I26754 Sequence 54
	21	23.8	0.5	82	4	AF294354S5	AF294358 Bos tauru
	22	23.8	0.5	87	14	AF050514	AF050514 Human end
	23	23.8	0.5	90	3	TRBANTATC2	M28495 Trypanosoma
	24	23.8	0.5	91	10	MMU130784	A1130784 Mus muscu
	25	23.8	0.5	94	6	AX326634	AX326634 Sequence
	26	23.8	0.5	94	6	AX326671	AX326671 Sequence
	27	23.8	0.5	94	9	S52152S19	S52191 CD11b-leuko
	28	23.8	0.5	96	5	AF025985	AF025985 Paratrich
	29	23.6	0.5	78	9	HSR305430	AF0305430 Homo sapi
	30	23.6	0.5	87	6	AR062856	AR062856 Sequence
	31	23.6	0.5	88	11	HUM07770A	L31186 Human STS U
	32	23.6	0.5	98	6	AX326603	AX326603 Sequence
	33	23.6	0.5	99	6	AX059602	AX059602 Sequence
	34	23.6	0.5	100	5	AF025983	AF025983 Paratrich
	35	23.4	0.5	77	6	A20465	A20465 Oligonucleo
	36	23.4	0.5	79	10	RATPLYC	M32437 Rat/polyoma
	37	23.4	0.5	83	6	AX240922	AX240922 Sequence
	38	23.4	0.5	88	9	S72771	S72771 Immunoglobul
	39	23.4	0.5	89	6	AX015186	AX015186 Sequence
	40	23.4	0.5	99	11	HSPE49C03	AL033823 H. sapiens
	41	23.2	0.5	87	5	CHRM1TAS	M19461 Chicken c-m
	42	23.2	0.5	87	5	OCU72578	U72578 Oryctolagus
	43	23.2	0.5	100	4	AF174523	AF174523 Bufo dana
	44	23	0.5	91	9	HSU2336	U32336 Human pre-B
	45	23	0.5	94	6	ARI70294	ARI70294 Sequence

ALIGNMENTS

RESULT 1
LOCUS DDIDDKD 87 bp DNA linear INV 27-APR-1993
DEFINITION D. discoideum protein kinase 4 gene, partial cds.

ACCESSION M59747.1 GI:167723
VERSION M59747.1
KEYWORDS protein kinase 4.
SOURCE Dictyostelium discoideum (strain AX-3) DNR.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
1 (bases 1 to 87)

REFERENCE Haribabu, B. and Dotin, R.P.
AUTHORS Identification of a protein kinase multigene family of Dictyostelium discoideum: Molecular cloning and expression of a cDNA encoding a developmentally regulated protein kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 88, 1115-1119 (1991)
MEDLINE 91142122
FEATURES location/Qualifiers
1..87
/organism="Dictyostelium discoideum"
/strain="AX-3"
/dbxref="taxon:44689"
!1..>87

RESULT
AX033159

LOCUS AX033159 70 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 3 from Patent WO0045176.
 ACCESSION AX033159
 VERSION AX033159.1 GI:10280023
 KEYWORDS
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE
 1 (bases 1 to 70)
 Gallusser, A., Karl, J., Lill, H., Stahl, P., Krueger, K. and Borgia, A.
 Method of identifying n-terminal propep
 Patent: WO 0045176-A 3 03-AUG-2000.
 GALJUSSE, ANDREAS (DE); KARL, JOHANN (DE); LILL, HELMUT (DE);
 STAHL, PETER (DE); KRUEGER, KERSTIN (DE); BORGIA, ANNELEISE (DE);
 ROCHE DIAGNOSTICS GMBH (DE)
 location/Qualifiers

FEATURES
 source 1..70
 /organism="Escherichia coli"
 /db_xref="taxon:562"

BASE COUNT 10 a 19 c 21 g 20 t

ORIGIN

Query Match 0.5%; Score 25.2; DB 6; Length 70;
 Best Local Similarity 62.9%; Pred. No. 2.7e+05;
 Matches 39; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

OY 1778 agggcgctctgtagaagcagctacagctcagctcagagagctgagagact 1837
 |||||
 Db 8 AGGAGGCTCTGTCAACCTGCACGCTTGGACAGTTTACCTCAGAGTGTGCTGTTCC 67
 |||||

OY 1838 tg 1839
 ||
 Db 68 TG 69

RESULT 6
 165769 99 bp DNA linear PAT 07-OCT-1997
 LOCUS 165769
 DEFINITION Sequence 5 from patent US 5668294.
 ACCESSION 165769
 VERSION 165769.1 GI:2482339
 KEYWORDS
 SOURCE unknown.
 ORGANISM unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 99)
 Meagher, R.B. and Summers, A.O.
 Metal resistance sequences and transgenic plants
 TITLE Patent: US 5668294-A 5 16-SEP-1997;
 JOURNAL Location/Qualifiers

FEATURES
 source 1..99
 /organism="unknown"

BASE COUNT 29 a 21 c 25 g 24 t

ORIGIN

Query Match 0.5%; Score 25.2; DB 6; Length 99;
 Best Local Similarity 54.3%; Pred. No. 2.8e+05;
 Matches 51; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

OY 3026 atgtcctgactaataatgtgagagtagaggtacaagaccaggtggcgaactcagat 3085
 |||||
 Db 2 AAGACCCAGCTATAGGTGAAGCTGTACTGCTTCATTTGCGATGGAAGCATGAACTGC 61
 |||||

OY 3086 ttgatttcaagaagaatgaactgccttat 3119
 |||||
 Db 62 GTGAGCATACTCAAGCAAGCAAGTGGCTATAT 95

RESULT 7
 RABPFKM18

LOCUS RABPFKM18 83 bp DNA linear MAM 27-APR-1993
 DEFINITION Rabbit muscle phosphofructokinase gene, exon 18.
 ACCESSION M14473 J02702
 VERSION M14473.1 GI:165622
 KEYWORDS phosphofructokinase.
 SEGMENT 18 OF 22
 SOURCE Rabbit (New Zealand) DNA, clone lambda-Charon 4APK.
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.

REFERENCE
 1 (bases 1 to 83)
 Lee, C.-P., Kao, M.-C., French, B.A., Putney, S.D. and Chang, S.H.
 The rabbit muscle phosphofructokinase gene: Implications for
 protein structure, function and tissue specificity
 J. Biol. Chem. 262, 4195-4199 (1987)
 MEDLINE 87166033
 COMMENT Draft entry and computer-readable sequence for [1] kindly provided
 by S.H.Chang, 02-FEB-1987.
 location/Qualifiers

FEATURES
 source 1..83
 /organism="Oryctolagus cuniculus"
 /db_xref="taxon:9986"

intron <1..15
 /note="pPK intron Q"

exon 16..77
 /number=18

intron 78..>83
 /note="pPK intron R"

BASE COUNT 23 a 13 c 27 g 20 t

ORIGIN About 719 bp after segment 17.

Query Match 0.5%; Score 25; DB 4; Length 83;
 Best Local Similarity 61.5%; Pred. No. 3.1e+05;
 Matches 40; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 1737 gagataagaagcgtgtagagagctgaagaggtctgaagggcgctctgataga 1796
 |||||
 Db 18 GAATGTGACACCTGTGGCAAGATGACGACAGCTGTGAAGAGCGTGTGGCTGAG 77
 |||||

OY 1797 gcaat 1801
 |||||
 Db 78 GTACT 82

RESULT 8
 AX173377 92 bp DNA linear PAT 03-JUL-2001
 LOCUS AX173377
 DEFINITION Sequence 31 from Patent WO0142445.
 ACCESSION AX173377
 VERSION AX173377.1 GI:14598152
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 Artificial sequence.
 REFERENCE 1 (bases 1 to 92)
 Murphy, B.R., Collins, P.L., Schmidt, A.C., Durbin, A.P.,
 Skladopoulos, M.H. and Tao, T.
 Skadiopoulos, M.H. and Tao, T.
 Use of recombinant parainfluenza viruses (pivs) as vectors to
 protect against infection and disease caused by piv and other human
 pathogens
 Patent: WO 0142445-A 31 14-JUN-2001;
 JOURNAL The Secretary of Health and Human Services (US)
 location/Qualifiers

FEATURES
 source 1..92
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Reverse primer for RSV A G gene insert"

BASE COUNT 20 a 12 c 26 g 34 t

ORIGIN

Query Match 0.5%; Score 24.8; DB 6; Length 92;

Best Local Similarity 63.3%; Pred. No. 3.5e+05;
Matches 38; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 45 taccctccccccagacacctgcagatgaagcagatgcacacacacacagagaa 104
Db 80 TTCTCTCCACCCACACACACGAGTAAATAAAGTAAAGAACTTAGATTA 21

RESULT 9
LOCUS ARI40801/c
DEFINITION Sequence 78 from patent US 6207816.
ACCESSION ARI40801
VERSION ARI40801.1 GI:14483297
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 71)
AUTHORS Gold, L., Janjic, N. and Pagratz, N.
TITLE High affinity oligonucleotide ligands to growth factors
JOURNAL Patent: US 6207816-A 78 27-MAR-2001.
FEATURES
source location/Qualifiers
1..71

BASE COUNT 15 a 16 c 22 g 18 t
ORIGIN

Query Match 0.5%; Score 24.6; DB 6; Length 71;
Best Local Similarity 65.5%; Pred. No. 3.9e+05;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgagatgagctgtattcaagcagatcacatgcagatcctcc 4012
Db 56 GAAGACGTAACTAGTATTAAACATACCCCTGCGCGATGCTCTCC 2

RESULT 10
LOCUS 193480/c
DEFINITION Sequence 78 from patent US 5731144.
ACCESSION 193480
VERSION 193480.1 GI:3937950
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 71)
AUTHORS Toothman, P.J., Ringquist, S. and Gold, L.
TITLE High affinity TGF-beta, nucleic acid ligands
JOURNAL Patent: US 5731144-A 78 24-MAR-1998;
FEATURES
source location/Qualifiers
1..71

BASE COUNT 15 a 16 c 22 g 18 t
ORIGIN

Query Match 0.5%; Score 24.6; DB 6; Length 71;
Best Local Similarity 65.5%; Pred. No. 3.9e+05;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgagatgagctgtattcaagcagatcacatgcagatcctcc 4012
Db 56 GAAGACGTAACTAGTATTAAACATACCCCTGCGCGATGCTCTCC 2

RESULT 11
LOCUS 195107/c
DEFINITION Sequence 78 from patent US 5731424.
ACCESSION 195107

VERSION 195107.1 GI:3939577
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 71)
AUTHORS Toothman, P.J., Ringquist, S. and Gold, L.
TITLE High affinity TGF-beta, nucleic acid ligands and inhibitors
JOURNAL Patent: US 5731424-A 78 24-MAR-1998;
FEATURES
source location/Qualifiers
1..71

BASE COUNT 15 a 16 c 22 g 18 t
ORIGIN

Query Match 0.5%; Score 24.6; DB 6; Length 71;
Best Local Similarity 65.5%; Pred. No. 3.9e+05;
Matches 36; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

OY 3958 gaacatgagatgagctgtattcaagcagatcacatgcagatcctcc 4012
Db 56 GAAGACGTAACTAGTATTAAACATACCCCTGCGCGATGCTCTCC 2

RESULT 12

LOCUS 576509/c
DEFINITION DIS8 (A) [hypervariable minisatellite M52] [human, individual
MACH, Genomic, 73 nt].
ACCESSION 576509
VERSION 576509.1 GI:243382

KEYWORDS
SOURCE human individual MACH.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 73)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

TITLE Monckton, D.G. and Jeffreys, A.J.
JOURNAL Minisatellite 'isolate' discrimination in pseudohomozygotes by
MEDLINE single molecule PCR and variant repeat mapping
92120671
REMARK Genbank staff at the National Library of Medicine created this
entry [NCBI gi243382 76509] from the original journal article.
This sequence comes from Figure 2.

FEATURES
source location/Qualifiers
1..73

gene /organism="Homo sapiens"
/db_xref="taxon:9606"
1..73
/gene="DIS8"
/allele="A"

BASE COUNT 41 a 0 c 0 g 32 t
ORIGIN

Query Match 0.5%; Score 24.4; DB 9; Length 73;
Best Local Similarity 60.6%; Pred. No. 4.5e+05;
Matches 40; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 4914 tgaatgcaaaagcgtattcgaatttaagaaaggctcttttcaataatgct 4973
Db 72 TATTTATTAATAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 13
OY 4974 ttattt 4979
Db 12 TTATTT 7

RESULT 13
LOCUS DROTRANSIN/c
DEFINITION DROTRANSIN
ACCESSION 98 bp DNA linear INV 05-JAN-1994

	DEFINITION	Drosophila melanogaster DNA sequence, complete insertion sequence.
	ACCESSION	L20950
	VERSION	L20950.1 GI:304832
	KEYWORDS	Insertion sequence.
	SOURCE	Drosophila melanogaster (Individual_Isolate B20.3) DNA.
	ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 98) Engels,W.R. and Johnson-Schultz,D.M. Deletion white allele. Transposase-induced derivative of white-hd8xk17 Unpublished (1993)
	REFERENCE	Johnson-Schultz,D.M. and Engels,W.R. F-element-induced interallelic gene conversion of insertions and deletions in Drosophila melanogaster Mol. Cell. Biol. 13, 7006-7018 (1993)
	JOURNAL	94019372
	MEDLINE	
	FEATURES	Location/Qualifiers source 1..98 /organism="Drosophila melanogaster" /isolate="B20.3" /db_xref="taxon:7227"
OY	BASE COUNT	37 a 2 c 1 g 58 t
	ORIGIN	
OY	Query Match	0.5%; Score 24.4; DB 3; Length 98;
	Best Local Similarity	56.1%; Pred. No. 4.5e+05;
	Matches	46; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
OY	1362	tctgacttcattatagaccatttggacacaagcactgaagaatggttaaga 1421
Dd	85	TATATAATATTATTAATTAATTAATTAATTAATTAATGAATTTATTTATAA 26
OY	1422	gttaattttgaagctcacaa 1443
Dd	25	TTTTATTTAATTAATATCATG 4
	RESULT 14	
	LOCUS	AR007408 100 bp DNA linear PAT 04-DEC-1998
	DEFINITION	Sequence 9 from patent US 5750497.
	ACCESSION	AR007408
	VERSION	AR007408.1 GI:396892
	KEYWORDS	.
	SOURCE	Unknown.
	ORGANISM	Unclassified.
	REFERENCE	1 (bases 1 to 100) Havelund,S., Halslr.o slashed.m.J., Jonassen,I., Andersen,A.Sloth. and Markussen,J. Acylated Insulin Patent: US 5750497-A 9 12-MAY-1998; Location/Qualifiers 1..100 /organism="unknown"
	TITLE	
	JOURNAL	
	FEATURES	
	source	
	BASE COUNT	29 a 23 c 22 g 26 t
	ORIGIN	
OY	Query Match	0.5%; Score 24.4; DB 6; Length 100;
	Best Local Similarity	54.4%; Pred. No. 4.4e+05;
	Matches	49; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
OY	4469	aagaataaacctcaagatagacttggaagtctctaactactcttatgtaa 4528
Dd	97	AAGGTCAACACCGTAAAGGTATCGTTGAACAATGTCTACTTCATCTGTTTGATC 38
OY	4529	tattacataaagactgctgagaagca 4558

```

Db          37  CAATTGAAAACTAGTGTGATTAGACGCAG 8
RESULT 15
LOCUS      U92173/c
DEFINITION Mus musculus clone 1C3 T cell receptor beta chain mRNA, partial
ACCESSION U92173
VERSION   U92173.1
KEYWORDS  GI:2894946
SOURCE    house mouse.
ORGANISM  Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
AUTHORS   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE      1 (bases 1 to 51)
JOURNAL    Johnston, S.L. and Wettstein, P.J.
MEDLINE    T cell receptor diversity in C3H.S specific for the CTT-1 and CTT-2
AUTHORS    minor histocompatibility antigens
TITLE      J. Immunol. 159 (6), 2606-2615 (1997)
JOURNAL    97444147
FEATURES   2 (bases 1 to 51)
SOURCE     Johnston, S.L. and Wettstein, P.J.
           Direct Submission
           Submitted (06-MAR-1997) Immunology, Mayo Clinic, 200 1st St S.W.,
           Rochester, MN 55905, USA
           Location/Qualifiers
             1..51
               /organism="Mus musculus"
               /strain="C57BL/6"
               /db_xref="taxon:10090"
               /clone="1C3"
               /cell_type="cytotoxic T lymphocyte"
               1..51
               /note="minor CTT-1 antigen specific CDR3 region"
               /codon_start=1
               /product="T cell receptor beta chain"
               /protein_id="AAC02878.1"
               /db_xref="GI:2894947"
               /translation="YCTSDPDVSGNTLYF"
BASE COUNT 12 a 13 c 11 g 15 t
ORIGIN
Query Match 0.5%; Score 24; DB 10; Length 51;
Best Local Similarity 75.0%; Pred. No. 5.6e+05;
Matches 30; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1513 gaattccagatccatgctggaggttcagatagtgacat 1552
| | | | | | | | | | | | | | | | | | | | | | | |
DB 41 GTATTTCACATACCTCGTGGGATACATCGACGAGTGCAGT 2

```

Wed May 29 16:25:19 2002

us-09-676-436-3.sz1ml100.rge

Page 6